UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

OFFICE OF
PREVENTION, PESTICIDES,
AND TOXIC SUBSTANCES

December 17, 2003

MEMORANDUM

SUBJECT: Transmittal of the Report and Meeting Minutes of the Endocrine Disruptor Methods Validation Subcommittee under the National Advisory Council for Environmental Policy and Technology (NACEPT) held June 5 - 6, 2003.

TO: Dorothy Bowers, Chair
National Advisory Council for Environmental Policy and Technology
Office of Cooperation and Environmental Management
And
Mark Joyce & Sonia Altieri
Designated Federal Officials
National Advisory Council for Environmental Policy and Technology
Office of Cooperation and Environmental Management

FROM: Jane Scott Smith, Designated Federal Official
Endocrine Disruptor Methods Validation Subcommittee
Office of Science Coordination and Policy, OPPTS

THRU: Joseph Merenda, Chair
Endocrine Disruptor Methods Validation Subcommittee
Director, Office of Science Coordination and Policy, EPA

Please find attached the minutes of the NACEPT Endocrine Disruptor Methods Validation Subcommittee Seventh open meeting held in Washington, D.C. from June 5-6, 2003. This meeting summary covers the status/results of the prevalidation work on the aromatase assay, steroidogenesis assay, the one generation extension study, and the mammalian two generation assay; and provided input and advice on the EDSP's validation plans for the steroidogenesis assay and mammalian two generation assay.

Information about this NACEPT EDMVS meetings and activities can be obtained from the website at http://www.epa.gov/scipoly/oscpendo or the OPPT Docket, Docket Number

OPPT-2003-0016 online (www.epa.gov/edocket) or at (202) 566-0280. Interested persons are invited to contact Jane Smith, EDMVS Designated Federal Official (DFO), via e-mail at smith.jane-scott@epa.gov.

cc: Charles Auer, OPPT
Daiva Balkus, OCEM
Sandra Evalenko, OPPTS
Elaine Frances, ORD
Susan Hazen, OPPTS
James Jones, OPP
Bill Jordan, OPP
Joseph Merenda, OSCP
Margaret Schneider, OPPT
Adam Sharp, OPPTS
OPPT Docket # OPPT-2003-0016

REPORT
OF
ENDOCRINE DISRUPTOR METHODS VALIDATION SUBCOMMITTEE MEETING
A Subcommittee of the National Advisory Council for
Environmental Policy and Technology
June 5 – 6, 2003
AT
RESOLVE, 1255 23RD STREET, N.W. SUITE 275
WASHINGTON, D.C.

This meeting was a review and discussion of the results and status of the prevalidation work on the aromatase assay, steroidogenesis assay, the one generation extension study and the mammalian two generation assay. Also to provide input and advice on the EDSP's validation plans for the steroidogenesis assay and mammalian two generation assay.

Jane Scott Smith, DFO	**Joseph Merenda, Chair**
Endocrine Disruptor Methods	**Endocrine Disruptor Methods**
Validation Subcommittee under	**Validation Subcommittee under**
The National Advisory Council for	**The National Advisory Council for**
Environmental Policy and Technology	**Environmental and Technology**
Date: _____	**Date: _____**

TABLE OF CONTENTS

EDMVS Members in Attendance at the June 2003 Meeting

Joseph Merenda, Chair
U.S. EPA

William Benson, Ph.D., Vice Chair
U.S. EPA

Theodora Colborn, Ph.D
World Wildlife Fund

Robert D. Combes, Ph.D.
Scientific Director, FAME

Peter L. deFur, Ph.D
Commonwealth University

J. Charles Eldridge
Wake Forest U. School of Medicine
Authority

David Hattan, Ph.D.
Food and Drug Administration

Robert J. Kavlock, Ph.D.
U.S. EPA (acting chair 7/23)

Timothy Kubiak, M.P.H.
U..S. Fish and Wildlife Service

Mildred Christian, Ph.D. (by Phone –Fri)
Argus International

James T. Stevens. Ph.D.
Wake Forest University

Ted Schettler, M.D., M.P.H.
Science & Environmental Health
Network

Gerald A. LeBlanc, Ph.D
North Carolina State University

Ron Miller, Ph.D
The Dow Chemical Company

Susan C. Nagel
U. Missouri - Columbia

James W. "Willie" Owens, Ph.D.
The Procter & Gamble Company

Thomas L. Potter, Ph.D.
USDA- Agriculture Research
Service

Penelope A. Fenner-Crisp, Ph.D.
ILSI Risk Science Institute

Glen Van Der Kraak, Ph.D.
University of Guelph

James D. Yager, Jr., Ph.D.
Johns Hopkins University

Paul Foster, Ph. D.
NIEHS

William Stokes, D.V.M.
NIEHS

Facilitator
Paul De Morgan
RESOLVE

Designated Federal Official
Jane Scott Smith
Office of Science Policy and Coordination

Presenters

In Order of Presentation

June 5, 2003

Jane Smith, DFO
EPA, OSCP

Dr. Jim Mathews
Research Triangle Institute (RTI)

Gary Timm & Dr. Les Touart
EPA, OSCP

Carol Sloan
Research Triangle Institute (RTI)

Gary Timm
EPA, OSCP

June 6, 2003

Dr. Paul Foster
NIH, NIEHS

Dr. Julia George
Research Triangle Institute (RTI)

Jim Kariya
EPA, OSCP

Oral Public Comment

In Order of Presentation:

June 5, 2003

Angelina Duggan, Ph.D.
CropLife America

Barbara Neal, Ph.D.
BBL Sciences

Rick Becker, Ph.D.
American Chemistry Council

Diana Zuckerman, Ph.D.
National Center for Policy Research for Women & Families

NOTICE

This meeting summary has been written as part of the activities of the National Advisory Council on Environmental Policy and Technology (NACEPT), Endocrine Disruptor Methods Validation Subcommittee (EDMVS). This meeting summary has not been reviewed for approval by the United States Environmental Protection Agency (Agency) and, hence, the contents of the meeting summary do not necessarily represent the views and policies of the Agency, nor of other agencies in the Executive Branch of the Federal government, nor does mention of trade names or commercial products constitute a recommendation for use.

The NACEPT EDMVS was established in partial fulfillment of a Congressional statute. When Congress amended the Federal Food Drug and Cosmetics Act (FFDCA) in the Food Quality Protection Act (FQPA) of 1996, it directed the U.S. Environmental Protection Agency (EPA) to develop a screening program to determine whether certain substances may have hormonal effects in humans. To ensure that EPA has the best and most up-to-date advice available regarding the validation of the screens and tests in the EDSP, EPA established the Endocrine Disruptor Methods Validation Subcommittee (EDMVS) under the NACEPT. The EDMVS provides independent advice and counsel to the Agency through NACEPT on scientific and technical issues related to validation of the EDSP Tier I and Tier II assays, including advice on methods for reducing animal use, refining procedures involving animals to make them less stressful, and replacing animals where scientifically appropriate. The EDMVS held their first meeting in October of 2001. This was the seventh meeting of the EDMVS.

The June 5 – 6, 2003 open meeting of the EDMVS was announced in the Federal Register on May 21, 2003 (Volume 68, Number 98). Further information about NACEPT EDMVS meetings and activities can be obtained from its website at http://www.epa.gov/scipoly/oscpendo or the OPPT Docket number OPPT-2003-0016 online at www.epa.gov/edocket or at (202) 566-0280. Interested persons are invited to contact Jane Smith, EDMVS Designated Federal Official (DFO), via e-mail at smith jane-scott@epa.gov.

National Advisory Council for Environmental Policy and Technology (NACEPT)
Endocrine Disruptor Methods Validation Subcommittee (EDMVS)
Plenary Meeting
June 5-6, 2003
Proposed Agenda

RESOLVE
1255 23rd Street, N.W., Suite 275
Washington, DC 20037
(202) 944-2300

Meeting Objectives:

- Review and discuss the status/results of the prevalidation work on the aromatase assay, steroidogenesis assay, the one generation extension study, and the mammalian two generation assay; and
- Provide input and advice on the EDSP's validation plans for the steroidogenesis assay and mammalian two generation assay.

Thursday, June 5, 2003

9:00 – 9:10 Welcome and Opening Comments
Joe Merenda, EDMVS Chair and Director, Office of Science Coordination and Policy (OSCP), EPA

9:10 – 9:30 Introduction, Agenda Review, and Review of Previous Meeting Summary
Paul De Morgan, Facilitator, RESOLVE

9:30 – 10:00 Review of EDMVS Work Plan
Jane Smith, EDMVS Designated Federal Official (DFO), OSCP, EPA

10:00 – 10:15 Review of EDMVS Member Selection Process
Jane Smith, EDMVS DFO, OSCP, EPA

10:15 – 10:30 Break

10:30 – 11:30 Presentation on Status of Aromatase Assay (Tier I): Optimization and Performance Comparison of the Assay Using Placental Tissues (Porcine, Human, Bovine) and Human Recombinant Receptor
Jim Matthews, Ph.D., Research Triangle Institute (RTI)

11:30 – 12:30 Update of OECD EDTA Activities
Gary Timm and Les Touart, Ph.D., OSCP, EPA

12:30 – 1:45 Lunch

1:45 – 2:30 Presentation and Discussion of Steroidogenesis Assay (Tier I): Preliminary Results of the Optimization of the Protocol using Sliced Testes
Carol Sloan, Ph.D., RTI

2:30 – 3:15 Presentation and Discussion of EDSP's Prevalidation and Validation Plans for Steroidogenesis
Gary Timm, OSCP, EPA

3:15 – 3:30 Break

3:30 – 4:30 Discussion of Steroidogenesis Plans and Related Issues

4:30 – 5:00 Public Comment
Members of the public will be given an opportunity to comment on any aspect of the EDMVS work. The amount of time given to each individual will depend on the number of people wishing to provide comment.

5:00 Adjourn

Friday, June 6, 2003

8:30 – 8:45 Settling In

8:45 – 10:30 Study Results, Comments, and Discussion of One Generation Extension
Julia George, Ph.D., RTI
Paul Foster, Ph.D., National Institute of Environmental Health Sciences (NIEHS)

10:30 – 10:45 Break

10:45 – 12:00 Presentation and Discussion of EDSP Validation Plan for the Mammalian Two Generation Assay
Jim Kariya, OSCP, EPA

12:00 – 1:15 Lunch

1:15 – 2:30 Continued Discussion of EDSP Validation Plan for Mammalian Two Generation Assay

2:30 – 3:00 Next Steps and Agenda for August Meeting

3:00 Adjourn

Introduction

The Office of Science Policy and Coordination's Endocrine Disruptor Screening Program established the Endocrine Disruptor Methods Validation Subcommittee (EDMVS) under The National Advisory Council for Environmental Policy and Technology (NACEPT). The first EDMVS meeting was held in October 2001. That initial meeting brought the members together to review the mission statement and discuss subcommittee roles and responsibilities. The second meeting, held in December 2001, was the first time the subcommittee members were presented with specific questions regarding assay protocols. This third meeting, held March 2002, continued discussions on protocols as well as some discussions on the validation process, Core Chemicals, 'low dose' and means of assessing human health effects. The fourth meeting, held as a teleconference, was wholly concerned with the Steroidogenesis assay. The fifth meeting held July 23-24, 2002, was concerned with screening criteria, core chemicals, In Vitro ER/AR assays, and dose setting as well as test results of two special studies, a pubertal study involving restricted feeding, and a mammalian 2-generation study involving PTU. Detailed review papers were presented on amphibian metamorphosis and invertebrate assays. The sixth meeting, held as a teleconference, was to receive comments and advice on the Fish Lifecycle DRP (Tier II).

This seventh meeting held June 5 – 6, 2003 reviewed and discussed prevalidation results for the steroidogenesis assay, aromatase assay and the mammalian two generation assay as well as the validation plans for each.

Endocrine Disruptor Methods Validation Subcommittee (EDMVS)
Seventh Meeting
June 5-6, 2003

Meeting Summary

Final

On June 5-6, 2003, the U.S. Environmental Protection Agency (EPA) convened the seventh meeting of the EDMVS. The meeting objectives included:

- Review and discuss the status/results of the prevalidation work on the aromatase assay, steroidogenesis assay, the one generation extension study, and the mammalian two generation assay; and
- Provide input and advice on the EDSP's validation plans for the steroidogenesis assay and mammalian two generation assay.

Copies of presentation slides and other materials distributed at the meeting may be obtained by contacting Jane Smith at smith.jane-scott@epa.gov or 202/564-8476. Many of the materials also are available on the EPA website at http://www.epa.gov/scipoly/oscpendo/edmvs.htm. EPA has established an administrative record for this meeting under docket control number OPPT-2003-0016. The docket is available for inspection in the TSCA Nonconfidential Information Center, 1201 Constitution Ave. N.W., Washington, DC or online at www.epa.gov/edocket. The center is open from noon to 4 p.m., Monday through Friday, excluding legal holidays. The telephone number of the center is (202) 566-0280.

Thursday, June 5, 2003

I. Welcome and Opening Comments

Joe Merenda, Director of the Office of Science Coordination and Policy (OSCP) and chair of the EDMVS, welcomed the EDMVS and members of the public. He conveyed that the Endocrine Disruptor Screening Program (EDSP) has many programmatic decisions to make on a tight timeline and is eager to get input and advice from the subcommittee.

II. Introductions, Agenda Review, and Review of Previous Meeting Summary

Paul De Morgan, senior mediator with RESOLVE, introduced himself and asked the EDMVS members to identify themselves and their organizations (see Attachment A, List of Participants).

Jane Smith, Designated Federal Official (DFO) for the EDMVS, explained that the meeting was being held in accordance with the Federal Advisory Committees Act (FACA) and all materials distributed would be available through the docket and on the website. She invited anyone experiencing problems with the website or other concerns to contact her.

Mr. De Morgan gave an overview of the materials distributed to the members and reviewed the meeting agenda. He noted that time was allotted for public comment at the end of the first day of

the meeting. Mr. De Morgan then reviewed the meeting ground rules.

Mr. De Morgan noted that the December 2002 meeting summary was in the members' notebook of materials. He asked members whether they had any final comments regarding this summary. As there were no comments, the summary will be considered final.

III. Review of EDMVS Work Plan

Ms. Smith reviewed the status of individual assays in tier 1 and tier 2 and how these assays fit into the EDMVS meeting schedule. (As indicated above, copies of slides from Ms. Smith's presentation, "The EDMVS' Draft Work Plan," may be obtained from the docket or EPA website.) She noted that the program is beginning to transition from prevalidation to validation.

Ms. Smith conveyed that, based on this work plan, the August 2003 meeting agenda will be very full and will include discussions on aspects of the fish screen, aromatase, steroidogenesis, and male and female pubertal assays, as well as a discussion on strain and species. Members requested that EPA add a discussion on diets and caging for in vivo studies to the August agenda. For the October timeframe, EPA projected that the avian DRP and in utero through lactation protocol demonstration results will be ready. Members asked EPA to send documents related to upcoming meeting agendas as soon as they are completed. Ms. Smith agreed to work with RESOLVE to distribute documents as early as possible and thanked the EDMVS for their time spent reading materials and preparing for meetings.

IV. Review of the Member Selection Process

EPA notified EDMVS in March regarding the upcoming subcommittee selection process. Ms. Smith further described the process, explaining that members' two-year term will end in October and they are all eligible to nominate themselves for the next subcommittee term. EPA will accept solicitations for nominations through June 30. Specific instructions are available in the May 30, 2003 Federal Register notice, and Jane encouraged members to submit a CV along with nominations. EPA hopes to announce the list of members in September.

Mr. Merenda strongly encouraged current EDMVS members to seriously consider continuing with the subcommittee. He emphasized that EPA would like continuity when the EDMVS is rechartered.

V. Presentation on Status of Aromatase Assay (Tier I): Optimization and Performance Comparison of the Assay Using Placental Tissues (Porcine, Human, Bovine) and Human Recombinant Receptor

Gary Timm, OSCP, EPA, reviewed past activities with the aromatase assay. EPA conducted a survey of various methods that have been used for in vitro methods and concluded they would pursue a placental microsomal assay and human recombinant assay. While the H295 cell-based assay was also promising, it remains in the research phase. EPA moved ahead with the placental and recombinant assays, and to respond to EDMVS concerns about using human tissue, investigated the use of porcine and bovine placenta. Research Triangle Institute (RTI) found

several difficulties in using animal placenta, from practical issues such as seasonal breeding and challenges in collecting the tissue at farms to low levels of aromatase in certain placentas. Due to these challenges, EPA decided, in consultation with RTI, to terminate work on animal placenta and continue work on human placenta and human recombinant assays. Mr. Timm introduced Dr. Jim Mathews, RTI, to give an update on the status of the aromatase assay, with EDMVS discussion to take place at the August meeting.

Dr. Mathews discussed progress on the aromatase prevalidation. (As indicated above, copies of slides from Dr. Mathews' presentation, "Pre-Validation of the Aromatase Assay Using Human and Bovine Placental, and Human Recombinant Microsomes," may be obtained from the docket or EPA website.) The goal of the prevalidation study was to identify optimal factors and conditions under which to conduct the aromatase assay. He noted he was particularly looking for input on how to use the optimized assay to determine the effect of selected substances on aromatase activity.

Dr. Mathews discussed the bovine and porcine placentas, including characteristics, collection, and preparation for the preoptimization study in the lab. Based on the results, RTI concluded that human placentas are preferable to bovine and porcine placentas due to easier collection, better known morphology, microsomal protein yield, and high activity levels. Dr. Mathews further recommended establishing standard operating procedures for handling potentially infectious materials, as well as obtaining information on screening for infectious diseases by donors where available.

Using human placental tissue and human recombinant aromatase, RTI will optimize conditions using a factorial design. To determine variability of the optimized assay, three technicians will independently conduct the assay on three separate days. Dr. Mathews reviewed the chemicals to be included. Dr. Mathews noted that the study is currently designed to determine IC_{50} and optimize substrate concentration for V_{max}. However, K_m is the substrate concentration at $0.5V_{max}$, and current perspectives in inhibition kinetics recommend determining K_i rather than IC_{50}. Thus, he asked the EDMVS for advice on whether they should change the experimental design to determine K_i.

Discussion

Some members commented on the source of human recombinant and human placental tissue. They voiced concern that there is only one source for recombinant and encouraged EPA to contact GenTest to ensure the supply is guaranteed and stable. A member questioned whether EPA's study will get adequate information on variation between human placentas if only one placenta is being used. The response was that there is information suggesting that as long as the placenta is from a non-smoker, and as long as we have good performance criteria, variability among human placentas should not be a problem. Another member asked whether there were issues in using human tissue in assays. Dr. Mathews explained that they went to an institutional review board, which had no concerns with using human placenta. The board characterized it as waste tissue. A member conveyed that, when making decisions about assay materials and their sources, EPA should consider implications of international acceptance and practical issues of obtaining materials in each laboratory.

Members commented on chemical selection. One member said that the selection of test chemicals does not take advantage of information available in the literature. For example, many chemicals EPA selected have activity on other enzymes in steroidogenesis as well as on aromatase. Using chemicals specific to aromatase should be part of the standardization protocol to confirm that activity is indeed due to aromatase. Dr. Mathews noted that chemical selection was an EPA policy decision. Mr. Kariya, OSCP, EPA, further explained that EPA had difficulty obtaining certain specific chemicals from pharmaceutical companies. Dr. Mathews suggested they could ask pharmaceutical companies for chemicals that they do not plan to use in drugs. Other members questioned why there were more negatives than positives in the proposed list. Mr. Timm explained that EPA had chosen more positives and removed some because of availability. It is possible to reduce the number of negatives as well. A member encouraged EPA to consider what constitutes a positive as testing moves forward.

A member inquired about the cell-based assay and whether it could be used to address aromatase activity. Mr. Timm responded that one of the advantages of the H295R cell-based assay is that it seems to take into account the entire steroidogenesis pathway from cholesterol to estradiol. Thus, it may be a tool to look at aromatase as well as other parts of the pathway. This assay could address inhibition as well as the up- and down-regulation of the gene, which is not addressed by the human placental assay. However, the cell-based assay is relatively early in the research and developmental stage. If it is promising, EPA hopes to move it into an optimization and validation program. Dr. Susan Laws, ORD, EPA added that, as aromatase is regulated differently in different mammalian tissues, another advantage of the H295R cell is that the cell line contains many of the promoters.

Dr. Mathews answered members' questions about background activity in the assay. He said that, in pre-optimization using porcine and bovine placentas, the positive/negative NADPH blanks were at 6% of level found with NADPH for the human recombinant assay. With the human placental assays, only about one-tenth the activity level was found. Levels will be tested again at the optimization stage.

A member told Dr. Mathews that when he had worked with tritium, non-specific ^{3}H exchange with water was a problem. Dr. Mathews responded that they found no nonspecific tritium exchange. He is not concerned about tritium exchange because the hydrogen that is labeled on the substrate is not very acidic.

In response to Dr. Mathews question about K_i and IC_{50}, members made the following comments:

- The design should be changed to find K_i and K_m.
- It should first be determined whether the chemical is an inhibitor. If it is, then it would be appropriate to determine K_i so that there's a basis for comparing chemicals.
- As aromatase is a screening assay, IC_{50} may be appropriate if the assay is run at a fixed concentration of substrate K_i can be determined if IC_{50} determined first and inhibition is observed.
- It may be best to determine K_i and K_m, for the record.

Other comments given by members included the following:

- RTI's results were consistent with work by an active research group in CalDavis working with bovine and porcine placentas. In porcine aromatase, this group has noted a fair diversity of activity among species and strains. As we look ahead, consider whether or not diversity of enzyme activity and end products is a factor in extending results to other species.
- EPA should consider measuring the end product, estrone.
- In response to a question about inducibility, genetic polymorphism, and difference between sexes, Dr. Mathews noted that Bob Bruggemeier of Ohio State said there are no known ones that give rise to catalytic differences. RTI is keeping the possibility in mind as they run the assay.
- In response to questions about the high concentration of propylene glycol, a member commented that, according to Dr. Bruggemeier, it is known to work and not interfere with the assay.
- Use solid phase extraction rather than dichloromethane, although it is recognized that DCM quenches the reaction.
- Aromatase is inducible. There might be a natural substrate which would change the response in vivo. Although this consideration does not affect the assay, it could affect extrapolation of assay results to the *in vivo* situation. The question of the protease inhibitor cocktail being needed came up and it seemed understood that as long as the rates were linear, it was not necessary.
- Need to specify incubation times precisely—down to the minute!
- The laboratory conducting the assay needs to show linearity and set criteria to optimize androstenedione concentration.
- A substance could induce CYP-19 in vivo. This may not affect the assay in vivo, but it could affect data extrapolation.

Mr. Timm explained that EPA will present the optimized protocol, data comparing the human recombinant and human placental assays, and the proposal for the inter-lab validation study at the August meeting (now scheduled for the December 2003 meeting). At that time, EPA will ask EDMVS whether both assays should be pursued. EDMVS will eventually look at the battery as a whole, and aromatase is one candidate among other options.

VI. Update of OECD EDTA Activities

Mr. Timm and Dr. Les Touart, OSCP, EPA, discussed recent OECD activities. (As indicated above, copies of slides from Mr. Timm's and Dr. Touart's presentation, "2003 OECD Activities," may be obtained from the docket or EPA website.) They discussed progress on endocrine disruption, including the Endocrine Disruptors Testing and Assessment (EDTA) Task Force and its three validation management groups (VMGs): ecotoxicity, mammalian, and in vitro/non-animal. EDTA's role is to plan and execute prevalidation and validation of endocrine test procedures, to oversee the development of test guidelines, and to provide review and quality control of documents.

Mr. Timm explained that member countries, including the United States, as well as stakeholders are involved in the EDTA and VMGs. The prevalidation and validation of test methods could be

led by OECD, an individual country, or a combination of OECD and a lead country. OECD could serve as a coordinator for guideline development or simply a facilitator of information exchange during the process. For most guidelines in which the OECD has an interest, the US will be the lead country, coordinating the technical work. Peer review will be conducted through the National Coordinator comment process.

As the EDTA discusses validation, EPA will keep the EDMVS informed and ask members for input for the US position. Mr. Timm reviewed the 2003 OECD meeting schedule as well as the activities of the non-animal, mammalian, and ecotoxicity VMGs. The non-animal agenda includes assays on receptor binding, aromatase, steroidogenesis, reporter gene/transcriptional activation, and in vitro cell/tissue, as well as a discussion of quantitative structure-activity relationships (QSARs). The mammalian agenda includes validation of the uterotrophic, Hershberger, and enhanced OECD 407 assays, as well as thyroid hormone testing and shared work on testing and assessment. The ecotoxicity agenda includes validation of the fish and amphibian screens, fish and avian testing, and invertebrates.

Following their presentation, Mr. Timm and Dr. Touart took questions from the EDMVS.

Discussion

Many members commented on the review process for OECD work. Comments included:

- EPA should determine how and when to input the US scientific community's comments into OECD's work. If there is disagreement on scientific points, it will drag out approval.
- There should be time for the scientific community to examine and respond to the OECD expert review so new methods do not become accepted guidelines before they are appropriately reviewed.
- The peer review needs to be independent, transparent, and should avoid region-specific solutions.
- As adoption of tests are a federal action, they are subject to the Endangered Species Act and will be subject to public comment. It would be beneficial to consider requirements ahead of time.
- EPA and ICCVAM have already established review processes; OECD does not need to "reinvent the wheel."
- There should be an open process like SAP and ICCVAM for peer review of uterotrophic and other OECD assays.

Mr. Timm responded to these concerns, noting that participants in the June OECD meeting will discuss criteria for what would be acceptable as a peer review. He clarified that all methods will have to go through a SAP review as well, though the SAP conducts a different level of review if it is a secondary review process. Mr. Timm did not know whether there is a process by which US stakeholders could have easy access but conveyed that any stakeholder can provide public comment on a SAP review. Further, regardless of OECD activities, the US will conduct an independent peer review of the entire battery after individual tests are validated.

A member stated that, while regulators in other countries understand aspects of the technical peer review, they are not always willing to give up their purview to conduct the review. There is a lack of experience with reviews conducted in the manner of SAPs, and there are logistical difficulties in coordinating reviews between countries. The uterotrophic review will be an opportunity to gather a scientific group and take them through the process.

Dr. Touart replied to a member's question about a master list of chemicals for validation, explaining that there is a list of chemicals being considered. They will look at the uterotrophic and Hershberger assays as a starting point to find common chemicals, but there will also be chemicals unique to specific assays. EPA is still looking for a compound active in fish but not animals, and vice versa. EPA will examine the list and update the EDMVS.

Mr. De Morgan acknowledged EDMVS members' outstanding questions about how the US model of independent scientific peer review interacts with OECD's validation activities. He noted that the EDMVS will look to EPA, as the OECD representative for the US, to think about questions raised by subcommittee members and decide what next steps to take in understanding concerns and possible solutions. (See Section X for further discussion of this topic).

VII. Presentation and Discussion of Steroidogenesis Assay (Tier I): Preliminary Results of the Optimization of the Protocol Using Sliced Testes

Carol Sloan, RTI, presented information on the background, design, and results of Phase I optimizations of the steroidogenesis assay. (As indicated above, copies of slides from Dr. Sloan's presentation, "Optimization of the Sliced Testis Steroidogenesis Assay," may be obtained from the docket or EPA website.) She outlined several reasons that the sliced testis assay was chosen to study steroidogenesis and described the two phases of the assay design and protocol. Dr. Sloan then discussed the validation of the lactate dehydrogenase (LDH) assay. Phase I optimization is complete, and phase II is currently being analyzed. Dr. Sloan highlighted the importance of a biological interpretation of the data. They will determine whether the test is both a good measure and is practical enough to be transferred to other laboratories.

Dr. Sloan made the following clarifications in response to EDMVS members' questions:

- The cytoxicity method should be specific for Leydig cells rather than general, though they have not yet determined a method.
- We are changing the thickness of testicular slices as we change the size of the fragment.
- Every lab should use the same RIA kit. There are many testosterone kits available, but the one we used seems to be adaptable to the rat.
- The atmospheric composition of 5% CO2 is typical in incubators. It also reflects normal atmospheric levels to which animals are exposed. We adjust pH levels after the CO_2 is added. The atmosphere is kept closed rather than maintaining the CO_2 level. The oxygen levels run out to 24 hours have shown good results.
- Animals were purchased specifically for this test. Some animals in the past have been on other experiments that required castration, such as the Hershberger, so animals could be

reused. It is not possible for animal suppliers to freeze testes for our use, because the cell culture would not work.

VIII. Presentation and Discussion of EDSP's Prevalidation and Validation Plans for Steroidogenesis

Mr. Timm presented information on the steroidogenesis prevalidation and validation study plans. (As indicated above, copies of slides from Mr. Timm's presentation, "Prevalidation and Validation Study Plan for Sliced Testes Assay," may be obtained from the docket or EPA website.) He emphasized that, in order to fund the study, EPA must conduct it before the end of the fiscal year in September. He also noted that many issues with the steroidogenesis study plan are cross-cutting with other validation programs. His presentation gave EDMVS members an opportunity to assess the assay's reliability and relevance for detecting compounds affecting steroidogenesis, specifically, interference with signal transduction, cholesterol transport, and the conversion of cholesterol to testosterone. The purpose of the prevalidation studies is to obtain initial information on protocol transferability and serve as a primary test of relevance. The optimization stage of prevalidation is complete, leaving the baseline study, pilot studies, and multichemical studies. Mr. Timm explained that the initial list of reference chemicals were selected for known mode of action, though this list was limited by availability of chemicals, as many pharmaceuticals are difficult to procure. As a result, many of the prevalidation chemicals will be duplicated in validation in order to cover modes of action. Mr. Timm then discussed the selection of laboratories, including the number of laboratories. Based on available data, six to ten laboratories will be appropriate to achieve high confidence in reliability. EPA will select six labs with eight replications for validation. Mr. Timm discussed data analysis strategies, intra- and inter-laboratory analyses, and measures of variation among laboratories. At the end of validation, each laboratory will report that the protocol was followed, difficulties in executing the studies, raw data, and a summary of data. This information will be compiled into the validation study report, including control charts and intra-laboratory statistics, to identify outlying laboratories and the nature of discrepancies.

Discussion

A member asked a clarification question about whether there would be a check between prevalidation and validation of the steroidogenesis assay. Mr. Timm said that there will be a check before EPA proceeds to validation, allowing for possible modifications. Given time constraints, it is unlikely EPA can bring prevalidation data to the EDMVS before proceeding to the validation stage.

A member inquired how EPA intends to assess cytotoxicity specificity to Leydig cells, expressing concern that a false positive in the steroidogenesis assay could lead to an inaccurate conclusion about maximum concentration. Cytotoxic concentrations might be easily reached in an *in vitro* test like this, but might be impossible to achieve *in vivo*. If Leydig cells are knocked out at a concentration that is not achievable *in vivo* you might infer, incorrectly, that a chemical inhibits aromatase at a level that it really does note. Dr. Ralph Cooper responded that EPA should go back and look into what positive chemicals are available that are specific to steroidogenesis and for which the mechanism of action and toxicity is known.

Members also gave the following comments in response to the two presentations:

- There may not be a need for that many laboratories and for this many chemicals, as some on the list have multiple mechanisms, particularly in in vitro systems.
- Before sending the assay to multiple labs EPA should be certain that the chemical selection will yield the cleanest and best information.
- A specific calcium channel blocker, for which the mode of action is known, would be a good chemical selection for blocking steroidogenesis.
- The parallel between steroid production and cell viability for the LDH assay is unclear. Some of the compounds chosen may not manifest in the time course of the experiment. While EDS is a very good Leydig cell toxicant, its timed course of action may not fit with this assay. EPA should consider a cost-benefit analysis of the LDH bioassay, because a chemical could cause high LDH activity and still not interfere with steroidogenesis.
- Certain concentrations of fenarimol, which is listed as an aromatase inhibitor and a negative, could inhibit all the P450s along the pathway and inhibit testosterone secretion, making it a positive.
- With regards to the RIA, EPA should consider that the 125 I, used to tag the testosterone, decays at a fairly rapid rate.
- EPA should determine and specify the minimum specific activity of assay materials to avoid situations in which laboratories use old or otherwise unacceptable reagents that yield poor results.
- Concern was expressed by one member that the scope of the validation effort may not be proportional to the importance of the assay. That is, this is a huge effort for a small assay.

Mr. Timm asked the following questions of the EDMVS:

1. Does the EDMVS agree with the stated objectives and data interpretation in the Validation Study Plan?
    - Determine relevance by testing known inhibitors of steroidogenesis and reliability by measuring variability in testosterone production
    - Assay will detect interference with key steps in the steroidogenic pathway, therefore interference with steroidogenesis will result in a decrease or increase in measured testosterone relative to controls

2. Does EDMVS agree with the structure of the prevalidation program?
    - Use of two laboratories in prevalidation
    - A Baseline study (conducted in triplicate) in all participating laboratories to ensure that all are equally capable of running the protocol.
    - A Pilot study (conducted in triplicate) with the positive control substance and a cell toxicant
    - Study chemicals (conducted in duplicate) with all known modes of steroidogenesis inhibiting action

3. Does the EDMVS agree with the structure of the validation program?
    - 6 laboratories
    - 2 replicates
    - A Baseline study
    - A Pilot study with positive control
    - A coded chemical study involving
        - A negative chemical
        - A Leydig cell toxicant
        - Three known inhibitors of steroidogenesis

4. Have we selected appropriate measures of reliability?
    - Coefficients of variation across studies (study SD/overall mean across studies
    - Ratio of between-study to within-study standard deviation (study SD/average standard error within studies)
    - Comparison of individual within laboratory standard deviations to aveage within laboratory standard deviation

5. Are the number of replicates taken over both prevalidation and validation sufficient to generate robust statistics?
    - **Baseline studies**
        - Preval: 2 labs x 3 replicates = 6
        - Validation: 6 labs x 2 replicates = 12

    - **Positive Control** (aminoglutethimide)
        - Preval pilot: 2 labs x 3 replicates = 6
        - Validation pilot: 6 labs x 2 replicates = 12
        - Preval high dose: 2 labs x 9 chems x 2 replicates = 36
        - Validation high dose: 6 labs x 5 chems x 2 reps = 60

        - **Preval/val chemicals**
          2 labs x 4 chems x 2 replicates = 16
          6 labs x 4 chems x 2 replicates = 48
        - **Preval chemicals**: 2 labs x 5 chems x 2 replicates = 20
        - **Validation chemicals:** 6 labs x 1 chem x 2 replicates = 12

    Total of 228 studies

6. Chemical distribution: Should dosing solutions be prepared centrally or on site?
    - **Central prep:** minimizes variability, same doses used in all labs, same solvent, but question of stability
    - **On site:** less transportation/stability problems, better test of real world implementation

7. Dose verification: Do doses need to be confirmed by analytical chemistry?
    - **Analyze all**
    - **Analyze sample**
    - **Rely on audit**: save samples and retrospectively analyze on a "for cause" basis

– **If yes, should they be shipped back to central facility for analysis or analyzed on site?**

8. Naïve labs/trained labs issue
    - To what extent should labs receive training in the conduct of the assay?
    - Should labs be required to demonstrate competence in running the assay (e.g., by running positive controls)?

Members considered and gave the following comments on questions five, six, seven, and eight:

Question 5

- The subcommittee generally agreed with the study design but asked that experimental results from prevalidation be used to determine the number of labs to be used in validation.

Question 6

- EPA should report to EMDVS about dose selection at the August meeting.
- EPA should ensure that target doses are realistic for feeding the animals and therefore legitimate in terms of triggering a positive result.

Question 7

- The cost of verifying all dose samples for chemical composition analysis would be very high and unnecessary at this time. There is no reason to suspect variation in the response between laboratories.
- Samples will be archived regardless of whether they are analyzed, so it is not necessary to do so now.
- It is acceptable to rely on an audit rather than analyzing all samples.
- EPA should consult the ICCVAM report and other materials on their web site that addresses these questions.
- Testing would be expensive and probably not appropriate, though it is advisable to conduct a storage stability study for some of the compounds, particularly for the more exotic pharmaceuticals, to determine whether they are stable in solution. This information would be useful for the audit.
- If EPA has stability data before starting on selected compounds, there is no justification to analyze all samples.

Question 8

- Training, such as video, in-person demonstrations, and other methods, is important to success of study.
- The lead laboratory should set performance criteria for running assays to judge whether labs are competent to conduct all study techniques. A very specialized protocol is unnecessary, but labs must be competent within reason.

Mr. Timm stated that EDMVS input from this discussion will be fed into a modification of the work assignment. EPA will move into Phase 2 and update the EDMVS on progress at the August meeting. As members did not specifically comment on questions one through five, EPA agreed to send out the questions with a deadline for member comments.

IX. Public Comment

At the conclusion of the day's deliberations, members of the public attending the meeting were given the opportunity to provide comments. Mr. De Morgan indicated that each person's comments would not be captured verbatim in the meeting summary, but rather just briefly summarized. He encouraged all to submit their comments in writing to Ms. Smith for inclusion in the EPA docket and posting on the website. A few of the people making comments presented slides. (As indicated above, copies of slides from Dr. Duggan's, Dr. Neal's, and Dr. Becker's presentations, as well as Dr. Zuckerman's written comments, may be obtained from the docket or EPA website.)

Angelina Duggan, CropLife America

Dr. Duggan discussed her organization's support of the two-generation rat reproductive toxicity assay as a Tier 2 test, as either a definitive test following Tier 1 or as a Tier 1 by-pass option. She emphasized the need to demonstrate the test in its entirety. Dr. Duggan also expressed concern that the test is complex and resource-intensive and stated that there is insufficient evidence to support the F1 extension solely for endocrine disruption effects.

Barbara Neal, BBL Sciences

Dr. Neal discussed areolae and nipple retention (A/N R) as an endpoint for possible future inclusion assays. She suggested that studying A/N R around post-natal day 13 may be useful as a tier trigger to focus additional attention on male reproductive tract evaluation and that this strategy should be further evaluated. However, Dr. Neal stated that there are insufficient data to conclude that adding A/N R to a two-generation study would improve the assay's sensitivity or change the no observed adverse effect level (NOAEL).

Rick Becker, American Chemistry Council (ACC)

Dr. Becker provided comments on EPA's proposed validation plan for the mammalian two-generation test. He stated that some of the proposed endpoints appear useful, such as additional thyroid endpoints in adult animals, while others appear redundant. Dr. Becker said that each additional endpoint should be evaluated to ensure it improves risk assessment and increases the assay's sensitivity, specificity, or reliability for detecting adverse effects.

Diana Zuckerman, National Center for Policy Research for Women & Families

Dr. Zuckerman explained her organization's concern with the early onset of puberty in girls, which can put children at higher risk for a variety of physical and psychological problems. She noted that exposure to chemicals is one explanation for this early development. Dr. Zuckerman urged the subcommittee to promote thorough and timely research to help prevent such effects in girls.

X. EDMVS and OECD Activities

Following public comment, Dr. Merenda conveyed to EDMVS members that EPA was interested in their comments regarding how the subcommittee's work fits with OECD activities. He encouraged members to discuss the issue briefly and to send further written comments if necessary.

Members gave the following comments:

- The primary issue is to allow input from US stakeholders into the OECD process.
- EPA could extend the opportunity for comment to this group and the public, advising on the US position. EPA could then bring written comments to the OECD when agendas are established.
- OECD is not a regulatory body. Rather, it develops test guidelines for 25 member countries. The regulating authority of each country must then decide whether to accept these guidelines. The main question is timing the input so it comes at the most logical and useful time for EPA. It is better to comment after a peer review commissioned by OECD.
- EPA should begin soliciting comments from stakeholders in a very broad, public, transparent fashion.
- If EPA wants EDMVS to comment on the uterotrophic or other assays, they should clarify the subcommittee's responsibility. It is unclear whether EDMVS members are meant to examine and endorse another peer review or to give their own individual scientific judgments.

Mr. Timm clarified that the appropriate time to submit comments is when the National Coordinator sends out the draft document for comment. Before that time, OECD's deliberations are confidential, making it very difficult for EPA to do more than give progress updates to the EDMVS. Access to documents and detailed discussions has not yet been possible.

Mr. De Morgan pointed out that some of the members' concerns are driven by a lack of information or a lack of clarity. He expressed that it would help the subcommittee to hear further from EPA regarding their view of how and when to integrate comments from the EDMVS and the broader stakeholder community.

Dr. Merenda reflected concerns he heard regarding how the EDMVS can comment on technical discussions at OECD that may lead to a validation exercise for an assay. Dr. Merenda reviewed that these discussions take place in VMGs and the EDTA, to which EPA sends a representative.

Dr. Merenda also reviewed the issue of how peer reviews are conducted, noting that the process will be discussed at the upcoming joint meeting. OECD members hold a variety of views about the peer review process. The US proposal calls for a transparent and publicly accessible process. Dr. Merenda said EPA will report back to the EDMVS on the outcome of the discussion at the joint meeting. The current pilot peer review process for the uterotrophic assay is not identical to the US process, as it is conducted through letters rather than public meetings. However, EPA

concluded that the process is adequate to achieve transparency and involves proper expertise. Dr. Merenda concluded by proposing that EPA create a brief summary of the process for OECD involvement and circulate it to the EDMVS.

Friday, June 6, 2003

XI. Study Results, Comments, and Discussion of One Generation Extension

Mr. Kariya reminded the subcommittee that the one-generation extension study was done largely in response to the EDSTAC recommendations to add a few endpoints to the two-generation assay to further explore the assay and alternatives to the two-generation assay. He asked the subcommittee to focus on the two-generation assay and where an extended one-generation assay might fit with the two-generation.

Rochelle Tyl, RTI, outlined the objectives and approach of the one generation extension study and summarized the results. (As indicated above, copies of Dr. Tyl's presentation, "One Generation Extension Study of Vinclozolin and Di-n-Butyl Phthalate Administered by Gavage on Gestational Day 6 to Postnatal Day 20 in CD (Sprague-Dawley) Rats," may be obtained from the docket or EPA website.) The objectives of the study were to determine: 1) whether some of the effects from perinatal exposure to Vinclozolin (VIN) or to Di-n-butyl phthalate (DBP) that can be easily detected after puberty are missed in weanling animals of the F1 generation; and 2) whether some of these effects occur at an incidence that would go undetected if only one male per litter is retained past puberty and examined at adulthood. Dr. Tyl emphasized that the study was hypothesis driven; it was not designed to be a Tier 1, Tier 1.5, or Tier 2 assay. The two hypotheses tested were as follows:

- The "Standard Two-Generation Protocol" cursory examination of up to three F1 males per litter at weaning and only one F1 male at adulthood allows adverse reproductive effects that appear at and after puberty to be missed.
- Examination of three or more F1 males at or after puberty, in addition to the F1 males examined at weaning, will detect additional reproductive effects and provide a more complete and accurate characterization of the effects of the test compound.

Dr. Tyl noted that the study did not include histopathology or andrology. She outlined the key elements of the study approach:

- Vinclozolin (VIN) and dibutyl phthalate (DBP), two known and well-characterized anti-androgens, were used, each at two doses.
- The high dose of each compound was a known effect level.
- The low dose of VIN was expected to produce hypospadias and vaginal pouches that would be hard to detect in weanlings, but easier to detect in adults.
- The low dose of DBP was the LOAEL (lowest observable adverse effect level) for this compound.
- These compounds and the selected doses were identified by basic research protocols, and were used to test this hypothesis in rats.

Dr. Tyl explained that one study question has not been addressed yet: would effects observed on PND 95 have been observed if we only examined one adult male per litter in each group? To determine this, the entire data set would be used to create a Monte Carlo-type simulation to randomly select three males per litter from each of the litters and one male per litter of the adults. The simulation results would be used to calculate the probability that effects would be detected by looking at only one adult male per litter.

After reviewing the specific study results, Dr. Tyl presented the following summary conclusions:

- Specific male offspring malformations were detected on PND 95 but not on PND 21:
  - prostate dorsal lobe abnormal/reduced in size (VIN, both doses; DBP, high dose)
  - prostate ventral lobe abnormal/reduced in size (both compounds, both doses)
  - epispadias (VIN, both doses)

- The incidence of specific male offspring malformations detected on PND 95 was higher than the incidence of the same malformation observed on PND 21:
  - agenesis of all or parts of the epididymis(des) (high dose of both VIN and DBP)
  - hypospadias (low dose VIN)
  - missing/reduced in size/abnormal seminal vesicles (high dose of both VIN and DBP)

- The effects of VIN on the incidence of hypospadias and ventral prostate agenesis were more obvious at PND 95 than at PND 21. This effect was more apparent at the low dose than at the high dose.
  - Hypospadias was observed in 9.7% versus 15.8% of the animals on PND 21 and 95, respectively.
  - High dose animals exhibited hypospadias at 80.0% versus 98.6% on PND 21 and 95, respectively.

- The effects of DBP (high dose) on the incidence of epididymal agenesis on PND 95 was approximately twice that observed on PND 21, and thus were more obvious on PND 95 than on PND 21.

- Adverse effects on the weights of some male reproductive tissues were more apparent at PND 95 than on PND 21:
  - adjusted right or left testis weight (high dose VIN)
  - absolute right cauda epididymis weight (low dose VIN)
  - adjusted right cauda epididymis weight (low dose VIN and DBP)
  - absolute LABC weight (low dose VIN), adjusted LABC weight (high dose VIN and DBP)
  - absolute and adjusted Cowper's gland weight (high dose VIN)

- Adverse reproductive system effects in toto (structural malformations and other abnormalities) of the low and high doses of VIN and the high dose of DBP on F1 adult male offspring would most likely be statistically significant with either one or three adult males/litter, and would have been detected with either study design.

- Adverse reproductive system structural effects in toto at the low dose of DBP on F1 adult male offspring were clearly biologically significant but not necessarily or likely statistically significant, with either one or three adult males/litter, and provide an example of effects that would not likely be detected with either study design.

- The more males examined per litter, the better the characterization of the litter as responding or not responding adversely to exposure, and the smaller the variance term for pooled litters within each treatment group. The enhanced sensitivity with more males examined per litter would increase the likelihood of detection of effects as statistically and biologically significant.

- Also, for effects with low incidence, such as in the low dose DBP group in this study, the risk with fewer males examined per litter is that the effect might be missed, i.e., the litter would be designated as not responding, on the basis of the one male examined, if that male did not exhibit the effect.

Following Dr. Tyl, Paul Foster, National Institute of Environmental Health Sciences (NIEHS), presented some additional comments on the study. (As indicated above, copies of Dr. Foster's presentation, "Extended One-Generation Study with Antiandrogens," may be obtained from the docket or EPA website.) He again noted that the study was hypothesis' driven and the choice of compounds and study design were based on the specific issues to be addressed. He commented that the study should provide valuable information in guiding necessary amendments in protocols for tier 2 testing, possibly tailored tier 2 testing.

Dr. Foster described the critical endpoints for Vinclozolin (hypospadias, prostrate agenesis/malformations, vaginal pouch) and for DBP (epididymal malformations, hypospadias, testicular effects, permanent changes in anogenital distance (AGD) and nipples). He noted that there has been some debate on whether changes in AGD and areolae/nipples are indicators of disturbance in androgen status or true malformations indicative of a rare but permanent structural change. He commented that recent data have indicated that these changes are likely to be permanent. He added, however, that a continuum exists with lower dose levels of weaker antiandrogens producing non-statistically significant (transient) changes in adults.

Noting that the enhanced weanling necropsy done in the study was a far more detailed examination than would normally happen in a multigeneration study, Dr. Foster outlined some of the study results. He also noted some study compromises and unresolved issues:

- Because no histopathology was done it was difficult to verify some of the milder abnormalities detected by gross examination, particularly for low dose DBP and controls.
- In regard to animal numbers/litter and analysis of malformations
  - the real comparison is with what is undertaken normally on a multigeneration assay;
  - a statistical comparison between results from one adult and three adults per litter is desirable;
  - statistical differences between specific malformations at both ages, plus a statistical analysis by sample size.

In closing Dr. Foster summarized several study conclusions:

- The dose levels were selected to ensure a response and not to determine a NOEL.
- Some specific male reproductive malformations were detected at PND 95 but not at PND 21.
- The incidence of specific malformations detected at PND 95 was greatly increased over the same malformation at PND 21 even though animal numbers were approximately equal.
- Adverse effects on the weights of some organs were more apparent at PND 95 than 21.
- Are the changes in some parameters noted at or before weaning permanent? (e.g., Does the lack of a permanent effect on AGD or nipples constitute an adverse response? Does this need to be statistically significant?)

Dr. Foster also echoed Dr. Tyl's last two conclusions regarding better characterization and enhanced sensitivity from examining more males per litter.

Discussion

Several members commented on the importance of doing the Monte Carlo simulation and analysis to determine what effects would be observed by examining only one adult male per litter. Mr. Kariya explained that the simulation was left out due to concerns about the amount of resources it would require at the time and because it was not considered necessary for this particular analysis. However, he noted that EPA is still considering setting up an analysis to answer the question.

A member complimented the study team on the execution of the study, commenting that it was a beautiful piece of work for characterizing responses with a large number of animals and endpoints. She noted that the study met its objectives but raised a question for the group to consider: is the goal to characterize everything that can happen or to try to find an effect level? She noted that if the goal is to try to find an effect level, anogenital distance will detect it. She also commented that observing a change is generally sufficient to show that an effect has occurred. She said the question to ask is not whether the change is a malformation or not, but rather whether the change really occurred and how to make regulatory decisions based on that information. She acknowledged that observing malformations was important for this study but commented that decisions should not be based on malformations alone but on the interrelationship of observations.

A member shared his own analysis of the data and some of the conclusions he drew from that analysis. He noted that anogenital distance and nipples provided a very early sign as to mechanism and can be observed fairly easily. He pointed out that effects on traditional tissue weights are observed at both PND 21 and 95, noting that if malformations are defined by size then they will more often be detected at PND 95. Encouraging members to consider whether the results represent an overemphasis on visual detection of malformations, he posed several questions to the group: how should such results be used in a regulatory context? Is it necessary to pick up malformations in five tissues versus two – will it make a real difference in how the risks are characterized or risk assessment is done? What information is adequate under what conditions and what is "icing on the cake"? Another member pointed out, however, that the doses

were selected to have an effect, and they may not have been low enough to determine what is necessary to be sensitive.

A member observed that the changes being considered are for a multigeneration reproduction study for all chemicals that will be tested in the future. He commented that it is crucial to ensure that whatever changes are made to the protocol add real value in terms of either increasing the sensitivity of the protocol or providing the ability to qualitatively detect effects that would not otherwise be detected. He added that the multigeneration protocol is one of the most important protocols to have internationally accepted for characterizing the risk of not just endocrine active materials, but all materials. Noting that a multigeneration reproductive assay may be conducted without any prior information on endocrine activity, he commented that the protocol must be as effective as possible without including unnecessary elements that may prompt a regulator to declare an existing study inadequate by comparison.

Several members agreed that so many endpoints are not necessary to establish an effect, but enough endpoints should be included to establish a pattern of effect. One member commented that even if it is difficult to show statistical significance, the fact that a host of responses do not occur at all in control animals but are present even at a low level in experimental animals, is biologically significant. He suggested that these effects could indicate a different pattern of response, and rather than disregard them we should challenge statisticians to help us discern what they mean.

A member observed that the movement is away from using just a no-effect level to do risk assessment and regulatory decision-making. For example, the Food Quality Protection Act requires consideration of common mode or mechanism of action, and other programs at EPA have begun to discuss having similar requirements. The member also noted an emerging preference for using benchmark doses, which encourage the generation of more and better data and modeling. She commented, however, that the question remains as to when one needs to generate these kinds of data.

A member reminded the group that these assays will form the basis for compliance with a number of legal requirements, including the Endangered Species Act, to protect a wide range of species. These assays, which are being designed principally with human health in mind, will also form the data base for protecting many other species for which a difference in anogenital distance is very important and may have major reproductive consequences. He cautioned against narrowing the assays to only those few endpoints that are known specifically to have a precise functional counterpart in humans.

A member commented on the value of including endpoints such as anogenital distance for their value as cross-over indicators. He offered the example of TCDD-related compounds, which are antiestrogenic and also produce anogenital distance effects but do not cause nipple retention.

A member referred the group to a conclusion of the International Life Sciences Institute (ILSI) regarding anogenital distance. In 1998 an ILSI work group acknowledged that anogenital distance is a very sensitive indicator of antiandrogenic activity but it should not necessarily be used as a surrogate marker for subsequent endpoints because the dose that affects the anogenital

distance is not a reliable predictor of the dose at which affects on these other endpoints will occur. Another member disagreed, commenting that anogenital distance is not only a sensitive marker but one that does not go away. A third member noted that much work has been done since 1998 that might cause a work group to draw different conclusions if the questions were revisited.

A member commented that it would not be practical to tailor tier 2 tests. He stressed the importance of having one protocol to use in all cases. He commented that if a tier 1 assay is positive and an adequate tier 2 test is available, it is not necessary to have another test with additional endpoints.

A member commented on the absence of histopathology in the one-generation extension study. He observed that although EPA guidelines require only ten animals per sex, the OECD guidelines require full histopathology. So histopathology is done routinely in laboratories doing studies for global purposes, and it is one of the more sensitive indexes of exposure characterizing adverse effects.

XII. **Presentation and Discussion of EDSP Validation Plan for the Mammalian Two Generation Assay**

Mr. Kariya presented the validation plan for the tier 2 mammalian two-generation assay. (As indicated above, copies of Mr. Kariya's presentation, "Validation Plan for the EPSP Mammalian Tier 2 Assay," may be obtained from the docket or EPA website.) Key points of the plan included the following:

- Accept the OPPTS Guideline for Reproductive Toxicity, with additional thyroid endpoints, as valid for EDSP tier 2 purposes.
- Include clarifications of certain procedures and endpoints already in the guideline as part of the EDSP assay.
- Ask for independent peer review of the validity of the guideline and clarifications that EDSP proposes.

Mr. Kariya reported that additional plans include: 1) encouraging one generation extension of F1 when indicated by tier 1, and in all tier 1 bypasses; and 2) developing additional information on the relevance and reliability of the one-generation extension, for possible future inclusion in the EPSP assay.

Mr. Kariya explained why the OPPTS Guideline for Reproductive Toxicity is accepted as valid for EDSP purposes. The guideline is generally accepted as valid for regulatory assessment for reproductive toxicity, and for the purposes of EDSP, the official language states that reproductive toxicity can be "an effect produced by a naturally occurring estrogen" and can be an indicator of endocrine effects.

Mr. Kariya commented that the suggestion to consider additional endpoints came from the EDSTAC and that the Standardization and Validation Task Force (SVTF) suggested focusing on thyroid-specific endpoints because including all additional endpoints would be impractical. Two

EDMVS members pointed out that though the EDSTAC discussed these issues they did not reach agreement on them. Mr. Kariya added that the relevance of most of the thyroid-specific endpoints was shown in the PTU study, which the EDMVS reviewed.

Mr. Kariya outlined some concerns about declaring the two generation assay valid at this point:

- strain/species issues, about which EDSP is preparing a white paper
- interlaboratory variability
- thyroid: other mechanisms, sensitivity
- extension of F1: additional studies
- timing, if other studies are needed

Mr. Kariya then presented several questions for the EDMVS to discuss. Several members indicated they would provide written comments to EPA in addition to the comments summarized below.

Question 1: Does EDMVS agree that the additional endpoints/clarifications proposed for the two-generation assay are well-characterized and that further validation of this set of endpoints for use in EDSP tier 2 is unnecessary?

Individual members shared several specific comments, including the following:

- Whole-mount histology of mammary tissue of males is not necessary; mere observation of the nipples should be sufficient.
- Whole-mount histology of mammary tissue of males is acceptable as a triggered measure but not as a standard requirement.
- It may not be necessary to measure anogenital distance for both the F1 and the F2 generation.
- There is an advantage specifically to looking at the nipples specifically on day 13.
- There is value in including the thyroid hormones and thyroid weight, not the histology; the current OECD guidelines at least include the thyroid weight.
- Points such as testes location at necropsy likely will be covered in the current guidelines as they are written, so clarification here may not matter. Clarification on malformation and agenesis also is probably not necessary.
- The number of days until occurrence of the vaginal plug seems a relatively simple measurement to include.
- There may be value to weighing the prostates separately, particularly since the dorsolateral has more credence and importance for human risk assessment.

Members also discussed thyroid measures. One member commented that the proposed measures can be done fairly well without a lot of extra training. Another noted that the T3 did not prove to be useful. A member commented that thyroid normalization should focus on just the males because of the large degree of variability in females with cycles. Another member noted, however, that the females are not cycling at PND 21, so there should be less variability.

A member expressed concern that thyroid function and its outcomes are being bypassed because of analytical problems. Observing that one reason for measuring maternal thyroid is that it is very important for pup development, he commented that if there are problems assessing thyroid,

another method should be found. He observed that the current thinking seems to be that measuring TSH, T4, and pup brain weight is all that can be done at this point in terms of thyroid function and its impact on the developing nervous system in the pup. Noting that these measures seem very blunt, he suggested that the group should revisit the other ideas generated by the EDSTAC. Another member reported that the Center for the Evaluation of Risks to Human Reproduction and NIEHS recently held a workshop on thyroid endpoints, looking for new or better ways to examine thyroid function. He said that the participants were not able to identify any better measures; there are no generically available tools for looking at thyroid functions. Mr. Timm commented that EPA will prepare a detailed review paper on thyroid, covering in vitro and in vivo methods, to be available in December. The paper may also address some of the other techniques that EPA is not focusing on now as well as possible chemical challenges to validate assays.

Question 2: Does the EDMVS agree that the endpoints in the tier 2 assay will allow a compound to be identified as possibly having "an effect in humans that is similar to an effect produced by a naturally occurring estrogen" (or androgen/antiandrogen or thyroid mimic/inhibitor) in the absence of tier 1 data? If not, what other endpoints should be included, or what supplemental testing would be appropriate?

A member commented that the multi-generation study is capable of picking up adverse effects that would result from estrogen, androgen, or thyroid mediated mechanism, assuming the thyroid endpoints are added to the assay. The assay picks up these mechanisms and others.

A member distinguished that yes, the assay may identify something that may have an effect, but additional studies may be required to confirm that it does have the effect. She noted that it should not be assumed that no one will ever be asked to do anything more than the definitive two-generation study. Another member commented that without the tier 1 mechanistic information, based on the results of the two-generation test one could say the effects are consistent with this chemical being an estrogen, but one could not say it is an estrogen.

Question 3: Does the EDMVS agree that the procedures and endpoints in table 2 of the presentation should be listed explicitly, to ensure adequate examination?

A member commented that observations of the gubernacular cords in the cranial suspensory ligaments are acceptable, but having to measure their lengths is problematic. The position of the ovaries and the testes is the key signal of whether there is an alteration in the lengths.

A member observed that the clarifications generally cover things that one should—not must—be looking for. He noted that most of the items are easily covered, but some are outside of what laboratories are usually trained to examine. Another member commented that not everything covered in the clarifications is already included in the current guidelines. He added that some of the measures might lead to further validation or standardization.

Question 4a: If the EDMVS advises EPA to validate additional endpoints, can the new endpoints be validated separately from endpoints already in the reproductive toxicity assay?

A member commented that he saw no reason why new endpoints could not be validated separately. Another member commented that at some point the entire protocol needs to be run with all the intended endpoints in order to test its practicality and to determine whether additional endpoints interfere with the ability to conduct the assay overall and the ability to gather information on the already required endpoints. A member commented that for testing the practicality of the full protocol, one laboratory would probably be sufficient if the endpoints were already validated for reliability.

A member commented that another necessary element is to explain: 1) why the new endpoints are relevant; 2) what the history of their use has been; and 3) why they are needed, given that some are redundant.

Question 4b: If the EDMVS advises EPA to validate additional endpoints, is it necessary to validate all new endpoints in a two generation study, or can relevance and reliability he established into shorter assay?

A member commented that it would probably not be necessary to do the F2 generation either.

Question 4c: If the EDMVS advises EPA to validate additional endpoints, how many laboratories should be required for interlaboratory comparability?

Bob Combes suggested that a statistical analysis similar to what was done for the steroidogenesis assay validation plan could be done for the two-generation validation plan as well.

Question 4d: If the EDMVS advises EPA to validate additional endpoints, how many chemicals promoted of endocrine activity should be tested in validation?

A member responded that using relatively few substances should be acceptable, but he noted that the test substances should include negative control substances with generalized toxicity.

A member offered that the choice of chemicals is more important than the number. He cautioned not to include a chemical sure to produce the desired result. Rather, at least one of the compounds should be a moderately weak compound to provide a legitimate regulatory test.

Noting that the two-generation assay is intended to broadly protect reproductive function and help with the development of the reproductive system, a member suggested that the focus should be on including broad, well-done, apical endpoints rather than on mechanistic issues.

Question 5: Does the EDMVS agree that the one-generation extension study shows increased sensitivity and provides greater precision in dose/response assessment, which will be of use in risk assessment, when the F1 animals are allowed to mature to PND 95 than when they are sacrificed at PND 21?

A member commented that one argument for extending the study to PND 95 may be the lower level of technical difficulty of making observations at PND 95 compared to PND 21, which may be particularly important when the tests are performed by a range of commercial laboratories.

A member commented that the study did not prove the value of making the proposed changes to the protocol. He commented that a direct comparison of the two protocols is really necessary to decide whether the existing protocol is adequate and whether an expanded protocol either increases sensitivity or has the ability to detect effects the original protocol would not detect. Another member answered "no" to question 5, arguing that anogenital distance, nipples, and other measurements provide sufficient indicators of effects at PND 21. Another member disagreed. She commented that the question cannot be answered with the two doses used, noting that the dose response curves in the study do appear to shift to the left based on nipple retention and areolae.

Several members commented that a statistical analysis comparing one-per-litter results with three-per-litter results is necessary to answer question 5. A member suggested that the analysis be done with a benchmark dose, if possible, to compare relative sensitivity. He also suggested using two doses and a control to avoid the issue of no-effect level and begin to address the question of how many endpoints are necessary to determine a pattern of response.

A member commented that he was not sure he could agree that the extension study shows increased sensitivity and provides greater precision in dose response assessment. He proposed, however, that these points may not be the issue. He observed that the extension study does provide redundancy in data points, allow for the identification of syndromes, and allow for looking at multiple effects at multiple time points, all of which are invaluable to provide confidence in the data for interpreting results and assessing risks.

Commenting on the practical aspects of the extension, a member observed that extending a study can introduce more variability, particularly with a measurement such as anogenital distance that relies on technician experience and does not have a standardized technique across laboratories. She agreed with the value of establishing a pattern of effects but cautioned that it would require a lot of training to do these types of studies.

Additional Points

A member asked whether behavioral assessment could be added to the assay. Dr. Cooper responded that the possibility is being explored. Dr. Tyl added that many laboratories have experience with behavioral tests. She noted, however, that it has not been shown whether behavioral tests are more accurate than measuring hormone levels; if hormone levels regularly pick up effects at lower levels, then behavioral tests may not be necessary.

A member requested that the subcommittee further discuss guidance for dose levels at a future meeting.

XIII. Next Steps and Agenda for August Meeting

Before the meeting adjourned on June 6, EPA staff presented a list summarizing the key points and potential action items they had drawn from the subcommittee's discussions. See Attachment B, Meeting Reflections.

The group discussed where to hold the next EDMVS meeting and ended up suggesting EPA try to hold the August meeting in Colorado. EPA agreed to explore the possibility. The group noted that while the meeting had originally been scheduled for August 18-21, given the interest in going to Colorado, this might require some minor shifting of dates within that week.

XIV. Closing Remarks

Mr. Merenda thanked the EDMVS members for the productive discussion and the depth and quality of their comments. He also thanked the public for attending and the speakers for their presentations.

Attachments: A. EPA Reflections
B. Supporting Materials for the EDMVS

Attachment A

EPA Reflections

Take Home Messages--Aromatase

- Consider a 2-stage assay
  - 1st stage to answer whether chemical is an inhibitor
  - 2nd stage to develop quantitative measure
- Investigate changing the experimental design to calculate Ki and km rather than IC50 since they are independent of substrate concentration.
- Ask pharmaceutical companies to sell us aromatase inhibitors that will not be commercialized.
- Ask GeneTest about their studies comparing placental assays and their recombinant assay system.

Take Home Messages--Steroidogenesis

- More thought needs to be given to chemical selection.
  - Find chemicals with only a single, well defined mode of action
  - Consider use of a Ca++ channel blocker as a steroidogenesis inhibitor
  - Fenarimol may inhibit all P450
- Revisit the cell toxicity issue
  - Should focus on Leydig cells.
  - Leydig cell toxicity may be a relatively slow event compared with enzyme inhibition Steroidogenesis (cont)
- Dose selection should be carefully considered. Discuss at August meeting.
- Check the ECVAM website for guidance on dose verification and analytical chemistry.
- E-mail questions.

Attachment B

Background Materials for the EDMVS
June 5 – 6, 2003 Meeting
Docket – OPPT-2003-0016
Website: http://www.epa.gov/scipoly/oscpendo/

1. **General Procedural**
   - Proposed Agenda
   - December 4, 2002 EDMVS Final Meeting Summary
   - EDMVS Work Plan, Revised

2. **Aromatase – Optimization and Performance Comparison of Assays Using Placental Tissues (Bovine, Porcine and Human) and Human Recombinant (Tier I)**
   - Pre-Validation Study Plan and Study Protocol for the Aromatase Assay
   - Pre-Optimization for Substrate Characterization for Bovine Placental Microsomes (Letter Report)
   - Pre-Optimization for Substrate Characterization for Procine Placental Microsomes (Preliminary Data Summary 05-08-03)
   - Pre-Optimization for Substrate Characterization for Human Recombinant and Human Placental Microsomes (Letter Report on Phase I)
   - Validation Study Plan

3. **Steroidogenesis – Results of Optimization of the Protocol using Sliced Testes (Tier I):**
   - Study Plan to optimize the Sliced Testis Steroidogenesis Assay – July 25, 2002
   - Results from the Optimization of the Sliced Testes Steroidogenesis Assay (Draft Letter Report) May 19, 2003
   - Validation Plan for Steroidogenesis
   - Steroidogenesis Questions

4. **One-Generation Extension Study Results (Tier II)**
   - One-Generation Extension Study Results

5. **Mammalian Two-Generation Assay Validation (Tier II)**
   - History, Plan, and Validation of the Mammalian Two-Generation Assay